Interhemispheric differences in seasonal cycles of tropospheric ozone in the marine boundary layer: observation - model comparisons

Interhemispheric differences in seasonal cycles of tropospheric ozone in the marine boundary layer: observation – model comparisons

Richard G. Derwent[a]*, David D. Parrish[b], Ian E. Galbally[c], David S. Stevenson[d], Ruth M. Doherty[d], Paul J. Young[e], Dudley E. Shallcross[f]

[a] *rdscientific, Newbury, Berkshire, UK*

[b] *NOAA ESRL Chemical Sciences Division, 325 Broadway R/CSD7, Boulder, Colorado, USA*

[c] *CSIRO Oceans and Atmosphere, PMB1, Aspendale, Victoria 3195, Australia*

[d] *School of GeoSciences, The University of Edinburgh, Edinburgh, UK*

[e] *Lancaster Environment Centre, Lancaster University, Lancaster, UK.*

[f] *Biogeochemistry Research Centre, School of Chemistry, University of Bristol, Bristol, UK.*

* Corresponding author.

E-mail address: r.derwent@btopenworld.com (R.G.Derwent)

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/2016JD024836

Abstract

Marine boundary layer ozone seasonal cycles have been quantified by fitting the sum of two sine-curves through monthly detrended observations taken at three stations: Mace Head, Ireland and Trinidad Head, California in the northern hemisphere and Cape Grim, Tasmania in the southern hemisphere. The parameters defining the sine-curve fits at these stations have been compared with those from a global Lagrangian chemistry-transport model (STOCHEM-CRI) and from fourteen ACCMIP chemistry-climate models. Most models substantially overestimated the long-term average ozone levels at Trinidad Head whilst they performed much better for Mace Head and Cape Grim. This led to an underestimation of the observed (North Atlantic inflow – North Pacific inflow) difference. The models generally under-predicted the magnitude of the fundamental term of the fitted seasonal cycle, most strongly at Cape Grim. The models more accurately reproduced the observed second harmonic terms compared to the fundamental terms at all stations. Significant correlations have been identified between the errors in the different models' estimates of the seasonal cycle parameters; these correlations may yield further insights into the causes of the model – measurement discrepancies.

Key Points

- Observed and modelled ozone seasonal cycles can be quantified by fitting sine-curves.
- Models tend to overestimate ozone in northern hemisphere mid-latitude marine boundary layers.
- Fundamental and second harmonic terms are not always well simulated by models.

Keywords:

Tropospheric ozone, seasonal cycles, ozone production, ozone sinks, interhemispheric differences.

1. Introduction

Ozone is widely recognised as an important air pollutant with widespread impacts on human health, crops and vegetation [Monks et al., 2015]. It is the focus of much policy-making activity, the aim of which is to reduce ozone exposures to meet air quality standards, guidelines or criteria by reducing emissions of its precursors, oxides of nitrogen (NO_x) and volatile organic compounds (VOCs). Although many policy questions can be answered using observational networks, models are important tools in the policy formulation process for ozone. Regional-scale chemistry transport or air quality models are in widespread use in the policy-making process to assess and promulgate strategies to achieve satisfactory air quality.

As the intensity and frequency of ozone episodes fall in both North America and Europe, there is an increasing focus by policy-makers on the intercontinental transport of ozone by policy-makers [HTAP, 2010; Clifton et al., 2014; Cooper et al., 2015; Doherty, 2015]. In Europe, there are concerns that the progress achieved by the reduction of regional-scale ozone levels has been offset by a growth in the hemispheric ozone levels [Collins et al., 2000] that has been attributed both to anthropogenic and natural (e.g. stratosphere-troposphere exchange, Neu et al. 2014) sources. Although episodic peak ozone levels monitored at the European stations with the highest mean ozone levels have declined markedly since 1980, these episodic peak levels at the stations with the lowest mean ozone have not [Derwent and Hjellbrekke, 2012]. This has been explained by the influence of the hemispheric scale transport of ozone. In North America, global scale chemistry models [Emery et al., 2012; Fiore et al., 2002; Zhang et al., 2011] are utilised to calculate hemispheric scale ozone concentrations transported into regions where exceedances of the United States National Ambient Air Quality Standards (NAAQS) are documented. This is the so-called 'Policy Relevant Background' [US EPA, 2014] or 'North American Background' and further details are given in Lefohn and Cooper, [2015].

Whilst the use of regional-scale air quality models in policy formulation is long-standing, the use of global models is relatively recent [HTAP, 2010]. If such models are to provide reliable future guidance for intercontinental policy formulation, then we must have confidence in their performance. Currently, this confidence is established through comparison with observations. However, it has not been possible, so far, to explain the origins of any shortcomings found other than to suspect the adequacy and completeness of any emission inventories employed, as well as chemical mechanisms, boundary layer mixing and convection, deposition, stratosphere-troposphere exchange and lightning [Wild, 2007]. It has not been possible either to reconcile good agreement in one part of the model against poor agreement found elsewhere in these complex models

Here we focus on the seasonal cycle in ozone in both models and observations and attempt a detailed examination of both, using techniques that are described in previous work [Parrish et al., 2016]. Our aim is to understand which observed features of the seasonal

cycle in ozone are faithfully reproduced by models and which features disagree. Our focus is on the marine boundary layer (MBL) because it receives relatively little in the way of emissions from human activities, because it is isolated from the rest of the troposphere and because it suffers much less from nocturnal depletion under shallow boundary layers. In this way, the process representation in chemistry transport models should be somewhat more reliable compared with that for continental areas. Furthermore, the marine environments upwind of North America and Europe have played an important role in the identification of the global rise in ozone baseline levels and the importance of intercontinental ozone transport [Parrish et al., 2016]. A potential difficulty, however, is the accurate representation in models of the entrainment of ozone-rich free tropospheric air into the MBL.

Parrish et al. [2016] compared observed seasonal cycles at eleven MBL sites with the results from three global chemistry-climate models (CCMs). They found similar seasonal cycles between sites within hemispheric scale regions. Here, we consider observations from only three sites that are representative of different hemispheric scale regions, chosen so that there are a pair of stations to reflect the gradient between the northern and southern hemispheres and a pair to reflect North Pacific inflow versus North Atlantic inflow. We compare results across a much larger number of models in order to obtain a more robust evaluation of the abilities of current models to correctly reproduce the seasonal cycle of ozone in the MBL. The models include a global Lagrangian chemistry-transport model STOCHEM-CRI [Derwent et al., 2015], which has increased chemical complexity to treat range of emitted hydrocarbons [Utembe et al., 2010], and the set of fourteen models that took part in the Atmospheric Chemistry Coupled Climate Model Intercomparison Project (ACCMIP) exercise [Young et al., 2013], see the Supplementary Information attached to this paper.

2. Techniques

In this study, seasonal cycles of ozone were defined by least squares fits of sine functions to observed or model monthly mean ozone mixing ratios, as follows:

$$y = Y_0 + A_1 \sin(\theta + \phi_1) + A_2 \sin(2\theta + \phi_2) \quad (1)$$

where Y_0 is the annual average ozone mixing ratio over the entire set of observations or model results, A_1 and A_2 are amplitudes, ϕ_1 and ϕ_2 are phase angles and θ is a time variable that spans one year's period in 2π radians. The second and third terms on the right hand side of equation (1) are the fundamental and second harmonic terms of the fitted ozone seasonal cycle. In previous work [Parrish et al., 2016], we have shown how the five parameters: Y_0, A_1, A_2, ϕ_1 and ϕ_2, represented all of the statistically significant information regarding the average seasonal cycle in the observations or model results. The observation – model comparisons in this study are based on the analysis of these five parameters.

Full details of the sources of the observations and model results, together with the estimation procedures are given in the Supplementary Information and only a brief summary is given here. Attention was focussed on three marine boundary layer (MBL) baseline stations that have relatively long measurement records: Mace Head, Ireland (1989 – 2014) and Trinidad Head, California, United States of America (1990 – 2010) in the northern hemisphere and Cape Grim, Tasmania (1982 - 2010) in the southern hemisphere. The ozone observations employed for Mace Head and Trinidad Head have been carefully filtered to remove local influences but retain baseline levels as described in Parrish et al., [2016]. In all cases, ozone concentrations are consistently expressed as mixing ratios in units of nmol ozone per mol air, referred to as ppb.

Model ozone seasonal cycles were taken from a global Lagrangian chemistry-transport model STOCHEM-CRI and from the set of fourteen chemistry-climate models that took part in the ACCMIP exercise. The model seasonal cycles were based on monthly mean ozone levels, including all hours of the day and night for the lowest model layers of the model grid cells containing the observing stations. The thickness of the lowest model layers and the dimensions of the grid squares containing the three MBL stations varied enormously and no attempt was made to harmonise these differences by interpolation. Uncertainty is introduced into the comparisons discussed below through the spatial mismatch between the observations made at a single point and the model calculations that were effectively an average over single grid cell in the model. This issue for the 3 MBL stations is discussed in some detail in Parrish et al., [2016]. Details of the fourteen models from ACCMIP are given elsewhere [Lamarque et al., 2013; Young et al., 2013].

3. Results

3.1 Observed ozone seasonal cycles

The seasonal cycles in ozone at the three selected MBL stations: Mace Head and Trinidad Head in the northern hemisphere and Cape Grim, Tasmania in the southern hemisphere are illustrated in Figure 1 for the observations and the sine-function fits. The annual average ozone mixing ratio for each station has been added to each fitted fundamental curve to facilitate comparison with the observations. The observations and sine-function fits overlap almost exactly at all sites. The fitted curves pass through the 2-sigma confidence limits for each monthly mean at all sites and the root-mean-square deviations between the fits and the monthly means are 0.7, 0.5 and 0.2 ppb for Mace Head, Trinidad Head and Cape Grim, respectively.

The observed and fitted seasonal cycles for the northern hemisphere Mace Head and Trinidad Head stations exhibited peaks in April and minima in July to August. The fundamental fits, in contrast, exhibited peaks in early March with minima six months later (September). The second harmonic fit showed two peaks, one in April and the other six months later (October). The observed and fitted seasonal cycles for the southern

hemisphere Cape Grim station peaked in August and showed a minimum in January. The fundamental fit for Cape Grim exhibited a peak in August. The second harmonic fit showed the same two peaks as in the northern hemisphere (April and October).

In this manner, it has been possible to quantify using five parameters: Y_0, A_1, ϕ_1, A_2 and ϕ_2, the observed seasonal cycles for the three MBL stations without loss of features and details or distortions. The fitted parameters and their confidence limits are presented in the Supplementary Information. Notably, all parameters derived here from the observations at the three stations are consistent with the values reported in Table 2 of Parrish et al., [2016]; comparison of future model results can simply use these tabulated values without directly accessing and analysing the monthly mean data themselves. An examination of model seasonal cycles now follows using the same sine-curve fitting procedures.

3.2 Comparison of the observed and modelled seasonal cycles

In this section, the fitted sine-curves to the seasonal cycles in the observations and models are compared using the five parameters: Y_0, A_1, ϕ_1, A_2 and ϕ_2, defined above in equation (1). The aim is to ascertain how well the models are able to quantify the seasonal cycles across the three MBL stations which have been chosen so that there are a pair of stations to reflect the gradient between the northern hemisphere versus the southern hemisphere and a pair to reflect North Pacific Ocean inflow versus North Atlantic Ocean inflow. Bar graphs of the observations and model ACCMIP ensemble mean (ENSEMBLE) results, with the STOCHEM-CRI and individual ACCMIP member results included as points are used to examine whether the models are able to account for the range in the observed seasonal cycle parameters between the three MBL stations. The detailed values of all parameters are tabulated in the Supplementary Information.

3.2.1 Long-term average ozone levels, Y_0

The long-term average ozone levels, Y_0, are compared at the three MBL stations in Figure 2. The observed values of Y_0 were, in ascending order: Cape Grim, 25.0 ± 0.2 ppb; Trinidad Head, 32.0 ± 0.7 ppb and Mace Head, 38.9 ± 0.4, where the quoted uncertainty ranges are 2 – σ or 95% confidence limits. All the models gave Cape Grim the lowest Y_0 but they disagreed about which station had the highest Y_0. The models typically show: Cape Grim < Mace Head < Trinidad Head. Further, Figure 2 shows that the models overestimate the observations for the Trinidad Head station by 2 to 19 ppb and that this discrepancy is the largest among the three stations.

The STOCHEM-CRI model calculated Y_0 values in the order: Cape Grim, 28.1 ± 0.8 ppb; Mace Head, 32.3 ± 2.1 ppb and Trinidad Head, 35.0 ± 1.8 ppb, respectively which gave the incorrect order for the North Pacific – inflow versus the North Atlantic – inflow stations. A similar problem was found for the ACCMIP models, with the predicted Y_0 values for the ENSEMBLE: Cape Grim, 24.4 ± 0.2 ppb: Mace Head, 40.8 ± 0.3 ppb and Trinidad Head, 41.1 ±

0.2 ppb. If account were taken of the 2 – σ error bars, then STOCHEM-CRI and the ACCMIP-ENSEMBLE both gave Mace Head minus – Trinidad Head differences that were statistically indistinguishable from zero. In contrast, the observed Mace Head – Trinidad Head differences were highly statistically significant, 6.8 ± 0.8 ppb. This difference was first recognised by Parrish et al. [2009] who reported 7 ± 2 ppb higher ozone in all seasons arriving at European baseline stations versus those in North America, in close agreement with the current estimate. Parrish et al. [2009] could not provide an explanation for this difference, and none of the models was able to reproduce it.

Using tagged tracers in the STOCHEM-CRI model [Derwent et al., 2015], the Y_0 value for ozone undergoing intercontinental transport to Trinidad Head was 24.8 ± 1.6 ppb compared with 35.0 ± 1.8 ppb for ozone from all sources. The corresponding STOCHEM-CRI Y_0 values for Mace Head were 29.7 ± 2.2 ppb and 32.3 ± 2.1 ppb, which suggested a smaller local ozone contribution than for Trinidad Head. This difference in behaviour suggested a larger local North American contribution at Trinidad Head compared with the local European contribution at Mace Head in the STOCHEM-CRI chemistry transport model.

Of the 14 ACCMIP models, only one model in addition to STOCHEM-CRI (see Figure 2) gave a Y_0 value for Trinidad Head that fell within ± 10% of the observations. The remaining 13 models gave substantially larger Y_0 values. This left the ENSEMBLE average substantially larger also. For Mace Head, the ACCMIP models performed much better such that several models and the ENSEMBLE gave results that fell within ± 10%. However, the Mace Head – Trinidad Head differences were found to lie within the range from -3.21 to +6.83 ppb, with the ENSEMBLE at -0.33 ppb. These should be compared with the observed difference of +6.8 ± 0.8 ppb. Only one of the ACCMIP models gave a reasonable estimate of the Mace Head – Trinidad Head difference of +6.83 ppb, but then only by overestimating both Y_0 values by substantial amounts, of the order of 7 ppb.

3.2.2 Amplitudes of the fundamental, A_1

The detailed values of A_1 are shown on the left side of Figure 3. The observed amplitudes of the fundamentals, A_1, were found to be: Mace Head, 5.6 ± 0.6 ppb; Trinidad Head, 5.7 ± 0.9 ppb and Cape Grim, 7.1 ± 0.2 ppb. The A_1 values for the two northern hemisphere stations were statistically indistinguishable, with that for the southern hemisphere station significantly greater. The observed order of the stations was therefore: Mace Head ≈ Trinidad Head < Cape Grim.

Model performance for this parameter was generally poor. STOCHEM-CRI gave reasonable estimates for A_1 at Mace Head, 4.7 ± 3.0 ppb and Cape Grim, 8.2 ± 1.2 ppb but that for Trinidad Head was too low by a wide margin, 2.8 ± 2.6 ppb. The results for this model were outside the ± 10% range and the stations were ordered differently than the observations: Trinidad Head < Mace Head < Cape Grim.

Agreement for the amplitudes from the ACCMIP models was also poor overall. The ENSEMBLE gave its lowest A_1 value for Cape Grim, 2.9 ± 0.2 ppb, next highest for Trinidad Head, 3.4 ± 0.2 ppb and highest for Mace Head, 4.1 ± 0.4 ppb. Again, the stations were ordered differently than the observations: Cape Grim < Trinidad Head < Mace Head. Of all the ACCMIP models, only one gave an A_1 value for Mace Head within ± 10% of the observed and one (though not the same model) for Trinidad Head. There were no ACCMIP predictions of A_1 within ± 10% of that observed for Cape Grim because generally all simulated amplitudes were gross underestimations, see Figure 3.

The northern Hemisphere – southern Hemisphere difference as indexed by the Mace Head – Cape Grim difference was found to be -1.5 ± 0.6 ppb in the observations. This difference was reported to be -3.5 ± 1.6 ppb in STOCHEM-CRI which, although it was of the correct sign, was found to be a substantially overestimated, as shown by the gradients in Figure 3. The ACCMIP ENSEMBLE gave a difference of +1.2 ± 0.4 ppb, which although approximately of the correct magnitude, had the incorrect sign.

STOCHEM-CRI and most ACCMIP models underestimated the A_1 values for the Trinidad Head station. When taken with the overestimation problem with the Y_0 values for the same station described in section 3.2.1 above, the underestimation problem with the A_1 values may point to a common issue across the models.

The ACCMIP models also performed poorly for Cape Grim but, in contrast, STOCHEM-CRI performed well at this station. The ACCMIP models significantly underestimated the strength of the fundamental amplitude of the seasonal cycle at Cape Grim compared with the observations. Whereas a marked seasonal cycle was observed, little was predicted. This may point to an underestimation by all models of photochemical ozone destruction or an overestimation of photochemical ozone production during the Austral summer.

3.2.3 Phase angles of the fundamental, ϕ_1

To facilitate comparison of the phase angles of the fundamentals of the observations and models between the northern and southern hemispheres, northern hemisphere ϕ_1 values have been shifted by $-\pi$ radians when plotted in Figure 4 and discussed below. With this adjustment, fundamental phase angles overlapped corresponding to a spring maxima in both hemispheres in both observations and models. ϕ_1 values were found to be less negative (peak earlier in the spring) for the observations at Cape Grim, -2.10 ± 0.03 radians, and more negative (peak later in the spring) at the two northern hemisphere stations: Trinidad Head, -2.66 ± 0.16 radians, and Mace Head, -2.63 ± 0.06 radians. The two northern hemisphere stations had ϕ_1 values that were statistically indistinguishable. The average observed difference in ϕ_1 between the hemispheres (-0.54 ± 0.16 radians or 31 ± 9 days) indicated that the maximum of the fundamental term occurred one month later in the northern hemisphere compared with the southern hemisphere. This difference may reflect the greater importance of photochemical ozone production in the northern hemisphere

where the large majority of man-made ozone precursors are emitted.

The ACCMIP ENSEMBLE values of ϕ_1 for the three stations gave the same overall pattern as the observations, see Figure 4, with Cape Grim as less negative and the two northern hemisphere stations as more negative, that is to say, earlier and later spring maxima, respectively. The observed phases of all stations were well reproduced by the ACCMIP ENSEMBLE but the models exhibited substantial variability that was larger in the northern hemisphere. One of the ACCMIP models exhibited a phase angle at Trinidad Head that was up to 2.5 radians (5 months) later in the year compared with the observations. The STOCHEM-CRI model did not accurately reproduce that phase at any station and was nearly completely out of phase at Mace Head.

3.2.4 Amplitudes of the second harmonics, A_2

The amplitudes of the second harmonics in the observations showed a regular progression across the three MBL stations (Figure 3): Cape Grim, 1.7 ± 0.2 ppb; Mace Head, 3.0 ± 0.6 ppb and Trinidad Head, 3.5 ± 0.9 ppb.

Figure 3 demonstrates that the models have significant skill in reproducing the broad spatial pattern in the observed second harmonic amplitudes. The amplitudes of the second harmonic in the STOCHEM-CRI model were found within ± 10% for Cape Grim and Mace Head but overestimated the observed value for Trinidad Head by 45%. The ACCMIP ENSEMBLE indicated a steeper gradient across the three MBL stations, underestimating the observed second harmonic amplitude at Cape Grim by more than a factor of two and overestimating it at Trinidad Head by 25%.

A substantial number of the ACCMIP models gave second harmonic amplitudes that lay within ± 10%. However, it was not always the same ACCMIP members that performed well at each station. There was significant variability in the ACCMIP results, such that the amplitudes spanned nearly a factor of three from 2.7 to 7.8 ppb at Trinidad Head, from 2.0 to 7.2 ppb at Mace Head and from 0.1 to 2.5 ppb at Cape Grim.

On the whole, the ACCMIP models reproduced the amplitude of the second harmonic better than that of the fundamental. Figure 3 shows closer accord of the absolute magnitudes for the ACCMIP ENSEMBLE for A_2 than for A_1 and the standard deviations of the ACCMIP members are smaller for A_2 than A_1.

3.2.5 Phase angles of the second harmonics, ϕ_2

The phase angles of the second harmonics are presented in Figure 4, noting that no shifting by -π radians has been applied to the northern hemisphere stations as with ϕ_1 . The observed order of the ϕ_2 values was: Mace Head ≈ Trinidad Head and both more negative than Cape Grim with the phase angles of the two northern hemisphere stations statistically indistinguishable: Mace Head, -2.4 ± 0.2 radians and Trinidad Head, -2.3 ± 0.3 radians. The

observed difference in phase angle between the northern hemisphere versus the southern hemisphere stations was small but statistically significant at -0.6 ± 0.3 radians (16 ± 8 days later maxima in the northern hemisphere).

The ACCMIP ENSEMBLE estimated ϕ_2 values at Mace Head and Cape Grim with pinpoint accuracy, within 0.07 radians or 2 days but was 0.37 ± 0.03 radians less negative or 11 ± 1 days earlier at Trinidad Head. The ACCMIP ENSEMBLE was thus able to accurately reproduce the observed difference in phase angle between Mace Head in the northern and Cape Grim in the southern hemisphere (-0.46 ± 0.3 radians). STOCHEM-CRI accurately reproduced the observed ϕ_2 values at Cape Grim but was 0.44 ± 0.26 radians or 13 ± 8 days early at Trinidad Head. However, this model was a long way out at Mace Head for this parameter.

There was a large range in the estimated second harmonic phase angles between the individual ACCMIP members, however. The northern hemisphere stations exhibited ϕ_2 values between -3.2 and -1.3 radians (23 days later to 30 days earlier) and the southern hemisphere values ranged between -3.2 and 1.2 radians (41 days later to 85 days earlier) relative to the observations.

In accord with the behaviour found for the amplitudes, the ACCMIP models reproduced the phase of the second harmonic better than that of the fundamental. Figure 4 shows good agreement on average for both phase angles but the standard deviations of the ACCMIP members are smaller for ϕ_1 than for ϕ_2. This is even more pronounced when the variability is considered in days compared to radians.

3.2.6 Variations in interhemispheric differences between the ACCMIP models

Direct examination of the interhemispheric differences can provide additional information that is not apparent from examination of model – observation differences at the separate stations. Figure 5 shows the interhemispheric differences for the five parameters discussed in the preceding sections. They are plotted between the northern and southern hemisphere (NS), calculated from Trinidad Head and Cape Grim parameters, and between the Atlantic and Pacific (AP) inflow stations, calculated from the Mace Head and Trinidad Head parameters. In Figure 5, the standard deviations of the results of the fourteen ACCMIP members are annotated. A striking feature of these standard deviations is that in most cases the standard deviations of the differences are smaller than the standard deviations at the individual stations annotated in Figures 2-4. Propagation of error considerations lead to the expectation of larger standard deviations for the differences if the model errors at the individual stations are uncorrelated. Therefore, this feature indicates that the model errors are significantly correlated between the stations.

The correlation of the Y_0 parameters derived from the separate ACCMIP members is examined in Figure 6a. The Y_0 values are highly correlated between Mace Head and Trinidad Head (r^2 = 0.74). This correlation indicates that, in addition to an overall bias between Mace

Head and Trinidad Head, each ACCMIP member tends to further overestimate Y_0 at both of these northern hemisphere mid-latitude stations by a similar amount, although the magnitude of this overestimation varies between the models. The correlation of the Y_0 values between Cape Grim and Trinidad Head is much poorer ($r^2 = 0.15$) with the ACCMIP models accurately reproducing the Cape Grim values, on average. A general positive bias of Chemistry-Climate Models for lower tropospheric ozone has been discussed [e.g. Lamarque et al., 2012; Naik et al., 2013; Parrish et al., 2014]; however, the correlations in Figure 6a suggest that such overestimates are found in all of the ACCMIP members at northern mid-latitudes, but are not a general global feature. Young et al. [2013] and Parrish et al., [2016] also found model overestimates at northern but not southern mid-latitudes. If the bias is indeed limited to northern mid-latitudes, this regional difference may help to diagnose the cause of the problem. In this case, the problem could arise from model treatment of anthropogenic emissions (which are concentrated at northern mid-latitudes) or model treatment of ozone deposition to continental surfaces (which are also concentrated in that region) or potential "dynamic" influences (weather patterns tend to be more complex and variable due to the more pronounced land-sea contrasts).

The correlations for all five parameters derived from the separate ACCMIP members are compared for NS and AP in Figure 6b and the correlation plots for amplitudes and phase angles are included in the Supplementary Material (Figures S1 and S2). The magnitudes of the fundamental of the seasonal cycles (A_1) are significantly correlated between Mace Head and Trinidad Head ($r^2 = 0.58$) and between Cape Grim and Trinidad Head ($r^2 = 0.48$). These correlations indicate that about one-half of the variance between the different models and between the models and the measurements is due to problems within each model that are common to all three sites, and the other half of the variance is due to model problems that differ between sites.

The errors for the ACCMIP members are significantly correlated for most of these parameters for both the NS and AP comparisons. These correlations all indicate that the ACCMIP models differ in important respects in their treatment of the processes that drive the ozone seasonal cycle throughout the troposphere, and the correlations can potentially provide diagnostic information regarding the causes of the errors within the individual ACCMIP models. One difficulty with such diagnosis is the limited precision possible for the determination of the seasonal cycle parameters with only a few years (5 to 10 years for the ACCMIP models and a single year for the STOCHEM-CRI model) of model simulations; the confidence limits of these parameter determinations (Tables S1 – S3) are of the order of the model – measurement differences, which limits our ability to interpret the present results. Future examination of the correlations with improved precision could provide useful guidance for model improvement.

4. Discussion and conclusions

To understand the seasonal cycle of ozone in the MBL, a simple conceptual model has been employed as formulated in our previous study [Parrish et al., 2016] in which ozone produced photochemically in the free troposphere or in the continental polluted boundary layer or injected from the stratosphere is entrained into or advected into the MBL. The late winter to early spring maximum and the corresponding late summer minimum is a reflection of the domination of the ozone seasonal cycle by net photochemical destruction in the MBL [Ayers et al., 1992; Oltmans and Levy, 1994]. Faster destruction in summer versus winter accounts for the summertime minimum and wintertime maximum. Consequently, ozone maximises in late winter to early spring and this is the main driver of the fundamental harmonic term seen in Figure 1 at the three chosen MBL stations. There may also be contributions from seasonal cycles in the entrainment of ozone-rich free tropospheric air into the MBL. The observed seasonal cycles are not pure sine curves and there is evidence for secondary maxima in late autumn and 'shoulders' during the late winter, see Figure 1. This behaviour is reflected in a large contribution from the second harmonic term as described by Parrish et al., [2016] who first recognised and quantified this term which they found to be a robust feature of observations and models for MBL stations. They argued that the second harmonic resulted from a second harmonic in the seasonal cycle of the photolysis rate coefficient of ozone which acts as the main photochemical destruction sink for ozone and provided a detailed discussion of this issue (see Section 4.3 of Parrish et al., [2016]).

In this study, we have extended the Parrish et al., [2016] work by analysing the seasonal cycles of 14 ACCMIP models rather than the three models of Parrish et al., [2016], together with the STOCHEM-CRI model at three MBL stations: Mace Head, Trinidad Head and Cape Grim. These three stations allowed us to focus on interhemispheric differences, that is to say, northern versus southern hemisphere and North Pacific versus North Atlantic. Our main finding was that we could accurately describe the seasonal cycles in the observations and all model results by fitting sine-curves and deriving five parameters: Y_0, A_1, ϕ_1, A_2 and ϕ_2 in equation (1). The fundamental term: $A_1 \sin(\theta + \phi_1)$, described the majority of the seasonal variations in both observed and modelled ozone. However, a second harmonic term of the form: $A_2 \sin(2\theta + \phi_2)$, was required to generate an accurate fit to all sets of observations and model results. Together, the five parameters provided a convenient means of accurately quantifying observed and model seasonal cycles.

Armed with this analytical tool, a systematic assessment was made of the seasonal cycles produced by STOCHEM-CRI and the 14 ACCMIP members and their ensemble mean (ENSEMBLE). Compared to the fundamental, all models more accurately reproduced the observed second harmonic terms. This accurate agreement both in amplitude and phase angle suggested that the second harmonic term arose from a cyclic phenomenon that was well simulated by all models. The cycle of the actinic flux and its control of the photochemical destruction of ozone, is a strong candidate to explain the second harmonic term, as argued by Parrish et al., [2016]. However, despite the general agreement found between the observed and model terms: A_2 and ϕ_2, there was a large amount of variability

between the results from the different ACCMIP members. The representation of the photochemical destruction sink for ozone should be straight-forwardly represented in the ACCMIP models and it is not at all clear why there should be such large variability. Further analysis was beyond the scope of this study.

Despite the large increase in chemical complexity in STOCHEM-CRI compared with the ACCMIP models, there did not appear to be much improvement in performance for the second harmonic parameters: A_2 and ϕ_2, versus observations. The STOCHEM-CRI and ACCMIP ENSEMBLE values for A_2 agreed reasonably closely, except for Cape Grim, and those for ϕ_2, except for Trinidad Head. The increase in chemical complexity in STOCHEM-CRI was entirely in the photochemical processes leading to ozone production. Because there was little or no increase in the complexity of the ozone destruction processes, it was reasonable that the second harmonic terms should be similar, assuming a similar level of treatment of the solar actinic fluxes. There was no significant improvement in the performance of STOCHEM-CRI with respect to A_1 and ϕ_1 compared to the ACCMIP models, for reasons which are not clear without further detailed information of the formulation of the ACCMIP models.

The model treatments of the fundamental terms: $A_1 \sin(\theta + \phi_1)$ and of the individual parameters: A_1 and ϕ_1, were in many cases in poor agreement with those of the observations. STOCHEM-CRI reproduced the observed fundamental amplitudes well at Mace Head and Cape Grim but underestimated them at Trinidad Head. The ACCMIP ENSEMBLE only performed well at Mace Head, in contrast, underestimating A_1 at the other stations. The ACCMIP ENSEMBLE performed well for ϕ_2 at all stations in contrast to STOCHEM-CRI which only performed well at Mace Head.

Further work is required to work through the model discrepancies found here to ascertain candidate explanations and identify needed improvements in tropospheric ozone models. We have identified those features of the model seasonal cycles that appear to be reasonably well described, namely the photochemical ozone sinks. Entrainment of free tropospheric air is expected to be an important factor controlling the concentrations and seasonality of MBL ozone; it will be important to investigate model treatment of this entrainment and degree of isolation of the MBL. We have also identified a particularly large extent of variability in the simulated ozone seasonal cycles between the ACCMIP members. Detailed analysis of ozone budget terms (including, for example, the possible importance of halogen chemistry) will be required over and above that already performed by Young et al., [2013] before the causes of the model variability and detailed discrepancies can be established. We have also identified significant correlations between the parameters derived from the individual ACCMIP models. For example, the models generally overestimate the long-term average ozone levels (Y_0) at northern mid-latitudes (but not in the southern hemisphere) and the overestimates of the different ACCMIP members correlate between Mace Head and Trinidad Head, possibly suggesting a model difficulty in

treating anthropogenic emissions or surface deposition, and that this difficulty varies between models. Model errors in A_1 correlate between all three stations, while model errors in A_2 and ϕ_1 correlate between Mace Head and Trinidad Head, but not between the northern and southern hemispheres. We suggest that future work further investigating these correlations in more detail may yield further insights into the causes of model-measurement discrepancies. Until these issues can be resolved, large uncertainties remain in tropospheric model simulations of ozone transported on intercontinental scales.

Acknowledgements

The authors are grateful to P.G. Simmonds and T.G. Spain for providing the Mace Head data and for A.J. Manning for sorting the Mace Head data into baseline and non-baseline observations. The data analysed here are available from the sources cited in the Supplementary Information.

D. Parrish acknowledges support from NOAA's Climate Program Office.

R. Derwent acknowledges support from the Department for Energy and Climate Change UK, under contract CESA 002 and from Department for Environment, Food and Rural Affairs UK for the development of STOCHEM and the CRI mechanism.

P. J. Young acknowledges support by an Early Career Integration Grant from Lancaster University.

References

Ayers, G.P., S.A. Penkett, R.W. Gillett, B. Bandy, I.E. Galbally, C.P. Meyer, C.M. Elsworth, S.T. Bentley, and B.W. Forgan (1992). Evidence for photochemical control of ozone concentrations in unpolluted marine air. Nature 360, 446-449.

Clifton, O.E., A.M. Fiore, G. Correa, L.W. Horowitz, and V. Naik (2014). Twenty-first century reversal of the surface ozone seasonal cycle over the north-eastern United States. Geophysical Research Letters 41, 7343-7350, doi:10.1012/2014GL061378.

Collins, W.J., D.S. Stevenson, C.E. Johnson, and R.G. Derwent (1997). Tropospheric ozone in a Global-Scale Three-Dimensional Lagrangian Model and its response to NO_x emission controls. Journal of Atmospheric Chemistry 26, 223-274.

Collins, W.J., D.S. Stevenson, C.E. Johnson, and R.G. Derwent (2000). The European regional ozone distribution and its links with the global scale for the years 1992 and 2015. Atmospheric Environment 34, 255-267.

Cooper, O.R., A.O. Langford, D.D. Parrish, and D.W. Fahey (2015) Challenges of a lowered U.S. ozone standard, Science, 348(6239), 1096-1097, doi:10.1126/science.aaa5748.

Derwent, R.G., and A.-G. Hjellbrekke (2012). Air pollution by ozone across Europe.

Handbook on Environmental Chemistry, 163. Springer-Verlag, Berlin.

Derwent, R.G., A.J. Manning, P.G. Simmonds, P.G., and T.G. Spain (2013). Analysis and interpretation of 25 years of ozone observations at the Mace Head Atmospheric Research Station on the Atlantic Ocean coast of Ireland from 1987 to 2012. Atmospheric Environment 80, 361-368.

Derwent, R.G., S.R. Utembe, M.E. Jenkin, and D.E., Shallcross (2015). Tropospheric ozone production regions and the intercontinental origins of surface ozone over Europe. Atmospheric Environment 112, 216-224.

Doherty, R.M. (2015). Ozone pollution from near and far. Nature Geoscience, News and Views, 10th August.

Emery, C., J. Jung, N. Downey, J. Johnson, M. Jimenez, G. Yarwood, and R. Morris (2012). Regional and global model estimates of policy relevant background ozone over the United States. Atmospheric Environment 47, 206-217.

Fiore, A.M., D.J. Jacob, I. Bey, R.M. Yantosca, B.D. Field, and A.C. Fusco (2002). Background ozone over the United States in summer: Origin, trend and contribution to pollution episodes. Journal of Geophysical Research 107, D4275, doi: 10.1029/2001JD000982.

HTAP (2010). Hemispheric transport of air pollution 2010. Part A.: ozone and particulate matter. Air Pollution Studies No. 17. United Nations, Geneva, Switzerland.

Jenkin, M.E., L.A. Watson, S.R. Utembe, and D.E. Shallcross (2008). A Common Representative Intermediate (CRI) mechanism for VOC degradation. Part-1: gas phase mechanism development. Atmospheric Environment 42, 7185-7195.

Lamarque, J. F., et al. (2012). CAM-chem: description and evaluation of interactive atmospheric chemistry in the Community Earth System Model. Geosci. Model Dev., 5, 369-411, doi:10.5194/gmd-5-369-2012.

Lamarque, J.F., et al. (2013). The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics. Geoscientific Model Development 6, 179-206. http://doi.org/10.5194/gmd-6-179-2013.

Lefohn, A.S., and O.R. Cooper (2015). Introduction to the special issue on observations and source attribution of ozone in rural regions of the western United States. Atmospheric Environment 109, 279-281.

Monks, P.S., et al. (2015). Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer. Atmospheric Chemistry and Physics 15, 8889-8973.

Naik, V., et al. (2013). Pre-industrial to present-day changes in tropospheric hydroxyl and methane lifetime from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP). Atmospheric Chemistry and Physics 13, 5277-5298.

Naik, V., L. W. Horowitz, A. M. Fiore, P. Ginoux, J. Mao, A. M. Aghedo, and H. Levy (2013). Impact of preindustrial to present-day changes in short-lived pollutant emissions on atmospheric composition and climate forcing. J. Geophys. Res., 118, doi:10.1002/jgrd.50608.

Neu et al., Tropospheric ozone variations governed by changes in stratospheric circulation Nature Geoscience 7, 340–344 (2014) doi:10.1038/ngeo2138.

Oltmans, S.J., and H. Levy (1994). Surface ozone measurements from a global network. Atmospheric Environment 28, 9-24.

Parrish, D.D., D.B. Miller, and A.H. Goldstein, (2009). Increasing ozone in marine boundary layer inflow at the west coasts of North America and Europe. Atmospheric Chemistry and Physics 9, 1303-1323.

Parrish, D.D., et al. (2014). Long-term changes in lower tropospheric baseline ozone concentrations: Comparing chemistry-climate models and observations at northern mid-latitudes. J. Geophys. Res., 119, doi:10.1002/ 2013JD021435.

Parrish, D.D., et al. (2016). Seasonal cycles of O_3 in the marine boundary layer: Observation and model simulation comparisons. Journal of Geophysical Research: Atmospheres 121, doi:10.1002/2015JD024101.

Sherrod, P.H. (1992). NLREG: Non-linear regression software. http://www.nlreg.com.

US EPA (2014). Policy assessment for the review of the ozone National Ambient Air Quality Standards. United States Environmental Protection Agency EPA-425/R-14-006, US EPA, Research Triangle Park, North Carolina.

Utembe, S.R., et al. (2010). Using a reduced Common Representative Intermediates (CRIv2-R5) mechanism to simulate tropospheric ozone in a 3-D Lagrangian chemistry transport model. Atmospheric Environment 44, 1609-1622.

Wild, O. (2007). Modelling the global tropospheric ozone budget: exploring the variability in current models. Atmospheric Chemistry and Physics 7, 2643-2660.

Young, P. J., et al. (2013). Pre-industrial to end 21st century projections of tropospheric ozone from the Atmospheric Chemistry and Climate Model Intercomparison Project. Atmospheric Chemistry and Physics 13, 2063-2090.

Zhang, L., et al. (2011). Improved estimates of the policy-relevant background ozone in the United States using the GEOS-Chem global model with 1/2^{o} x 2/3^{o} horizontal resolution over

North America. Atmospheric Environment 45, 6769-6776.

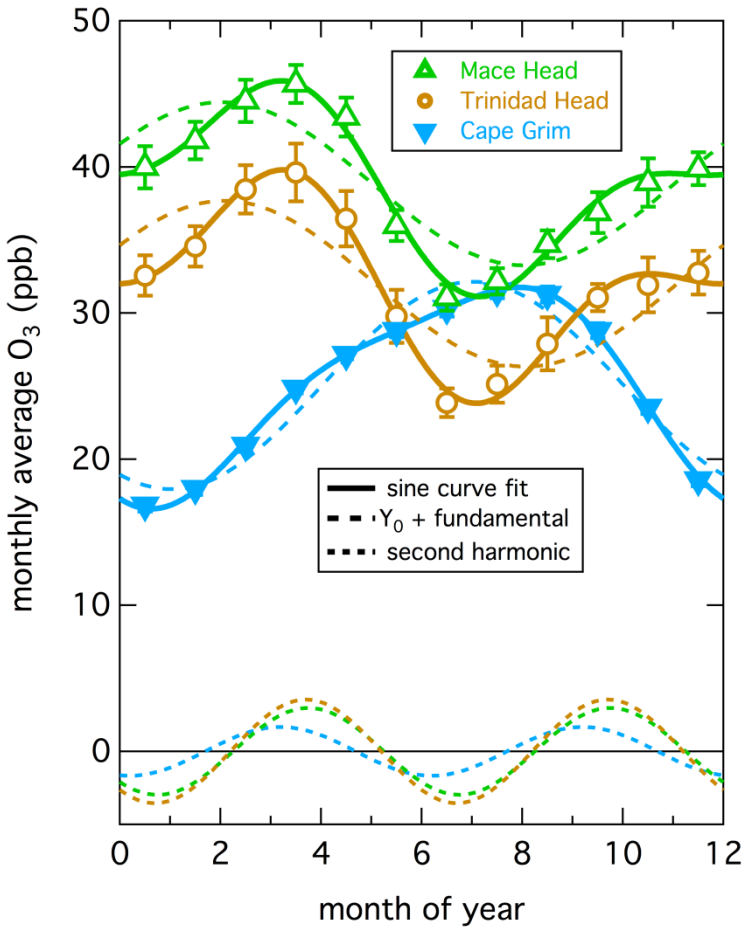

Figure 1. Sine-curve fitted ozone seasonal cycles (solid lines) to observations from Mace Head, Ireland (26 years), Trinidad Head, California (21 years) and Cape Grim, Tasmania (29 years), together with the fundamental and second harmonic fits (dashed lines). The symbols give average monthly ozone concentrations over the entire data records with error bars indicating 2-sigma confidence limits (some error bars are obscured by the size of the symbols).

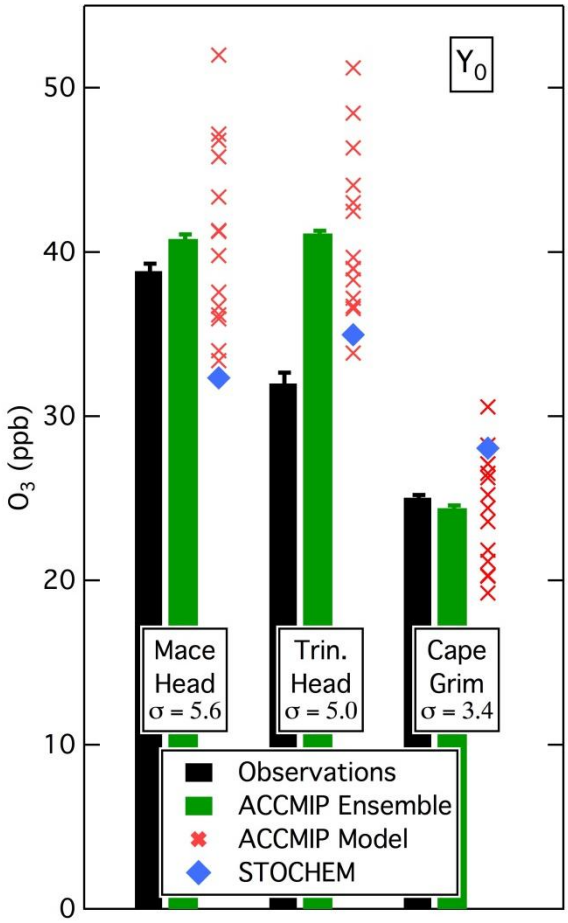

Figure 2. Comparison of the observed and model values of the average ozone mixing ratio over the entire dataset, Y_0, in ppb for the STOCHEM-CRI model, the ACCMIP members and their ENSEMBLE. Error bars indicate 2-sigma confidence limits for the observations and ACCMIP ensemble. The standard deviations of the results of the ACCMIP members are annotated for the three sites.

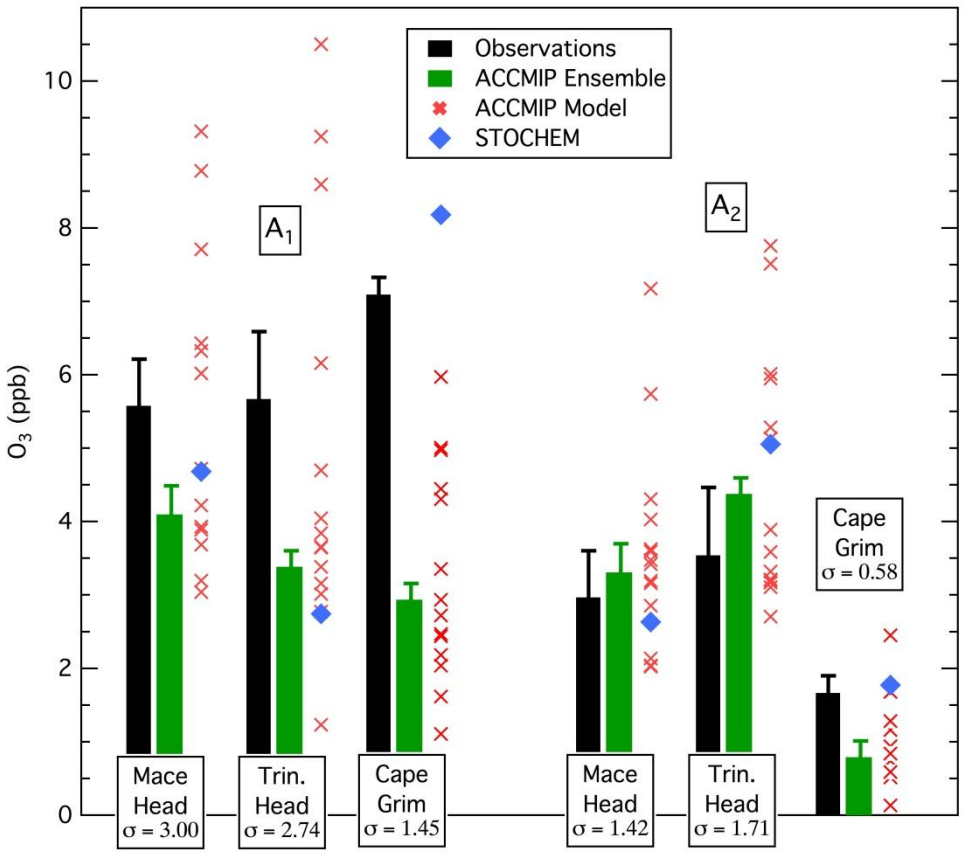

Figure 3. Comparison of the observed and model values of the fundamental and second harmonic amplitudes, A_1 and A_2, in ppb for the STOCHEM-CRI model, the ACCMIP members and their ENSEMBLE. Error bars indicate 2-sigma confidence limits for the observations and ACCMIP ENSEMBLE. The standard deviations of the results of the ACCMIP members are annotated for both amplitudes at the three sites.

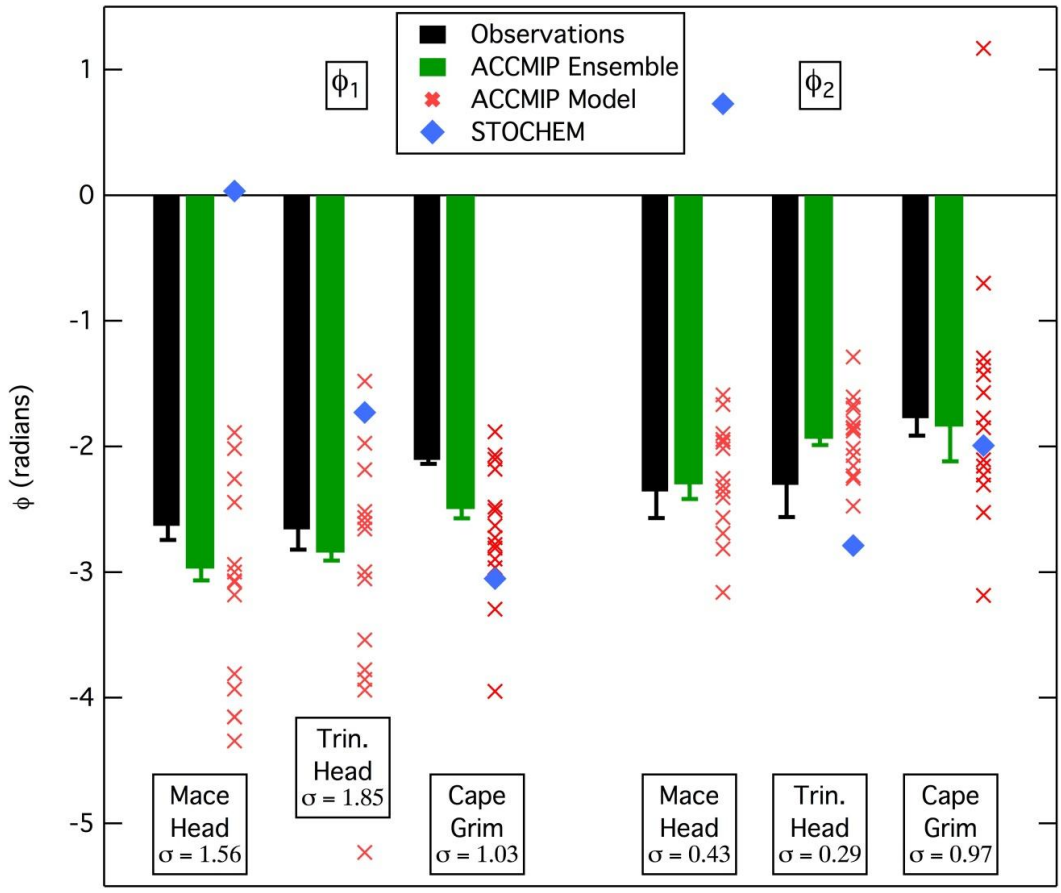

Figure 4. Comparison of the observed and model values of the fundamental and second harmonic phase angles, ϕ_1 and ϕ_2, in radians for the STOCHEM-CRI model, the ACCMIP members and their ENSEMBLE. The ϕ_1 values for the northern hemisphere stations have been shifted by -π radians to allow direct comparison with the southern hemisphere station. Error bars indicate 2-sigma confidence limits for the observations and ACCMIP ENSEMBLE. The standard deviations of the results of the ACCMIP members are annotated for both phase angles at the three stations.

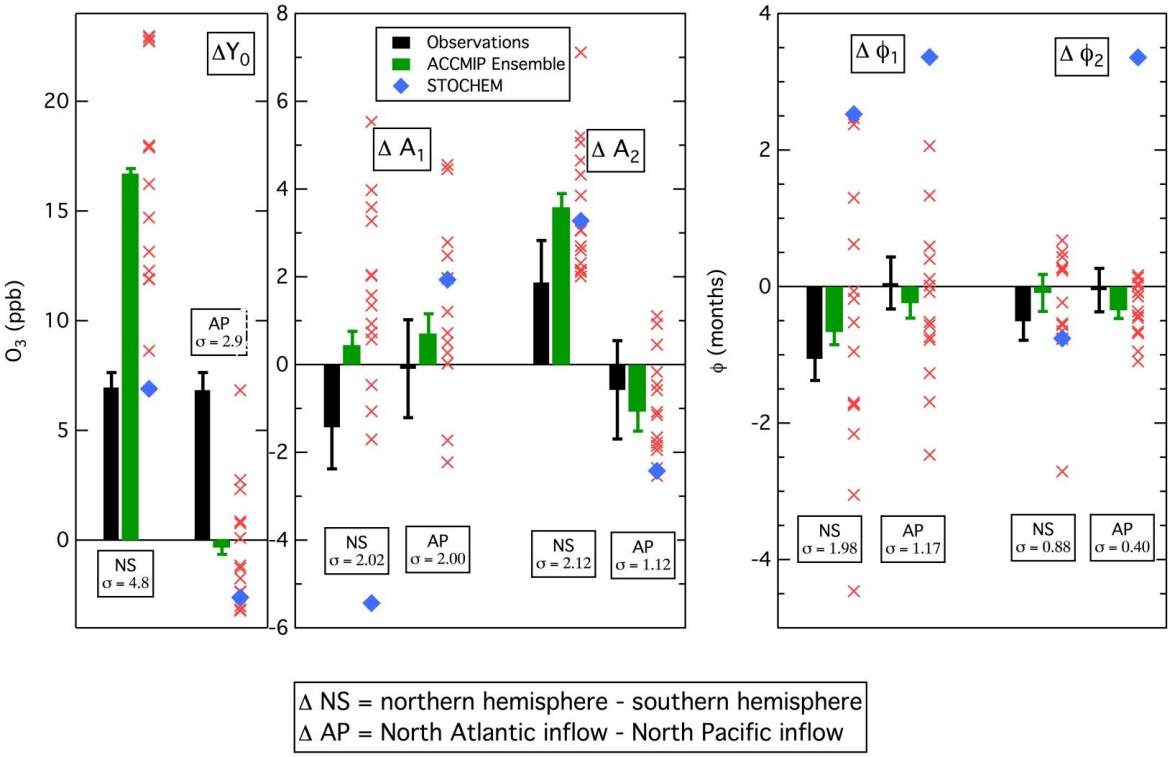

Figure 5. Comparison of the observed and model values of the differences in Y_0 (ppb), the two amplitudes (ppb) and the two phase angles (months), for the STOCHEM-CRI model, the ACCMIP members and their ENSEMBLE. The differences are between the northern and southern hemispheres (calculated from Trinidad Head - Cape Grim) with ϕ_1 at Trinidad Head shifted by π radians so that a difference of 0 indicates both have the same seasonal cycle, but shifted by 6 months between hemispheres) and between the Atlantic and Pacific inflow (calculated from Mace Head - Trinidad Head). Error bars indicate 2-sigma confidence limits for the observations and ACCMIP ENSEMBLE. The standard deviations of the results of the ACCMIP members are annotated for all parameters at the three sites.

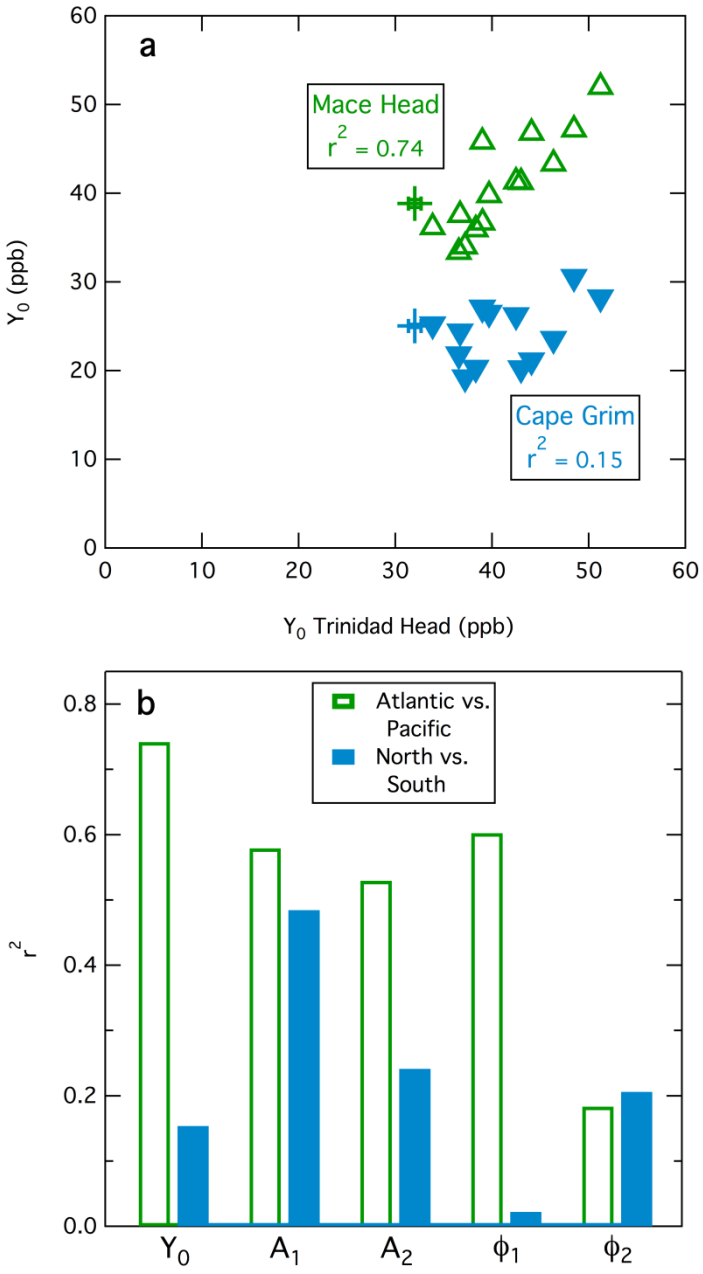

Figure 6. **a)** Correlation of Y_0 for Mace Head and Cape Grim with that for Trinidad Head. The triangles give the results derived by the fourteen ACCMIP members, and the plus symbols with error bars (smaller than the symbols) indicate the observations. The square of the linear correlation coefficient is annotated for each set of results. **b)** Comparison of the squares of the linear correlation coefficients for all five parameters derived from the fourteen ACCMIP members. The correlations are for North Atlantic versus North Pacific inflow (i.e., Mace Head versus Trinidad Head) and northern versus southern hemisphere (i.e., Cape Grim versus Trinidad Head).